

From regular to growing small-world networks

Zhongzhi Zhang,^{1,2,*} Shuigeng Zhou,^{1,2,†} and Zhen Shen^{1,2}

¹Department of Computer Science and Engineering, Fudan University, Shanghai 200433, China

²Shanghai Key Lab of Intelligent Information Processing, Fudan University, Shanghai 200433, China

(Dated: February 6, 2008)

We propose a growing model which interpolates between one-dimensional regular lattice and small-world networks. The model undergoes an interesting phase transition from large to small world. We investigate the structural properties by both theoretical predictions and numerical simulations. Our growing model is a complementarity for the famous static WS network model.

PACS numbers: 89.75.-k, 89.75.Fb, 05.10.-a

I. INTRODUCTION

Many real-life systems display both a high degree of local clustering and the small-world effect [1, 2, 3, 4, 5]. Local clustering characterizes the tendency of groups of nodes to be all connected to each other, while the small-world effect describes the property that any two nodes in the system can be connected by relatively short paths. Networks with these two characteristics are called small-world networks.

In the past few years, a number of models have been proposed to describe real-life small-world networks. The first and the most widely-studied model is the simple and attractive small-world network model of Watts and Strogatz (WS model) [6], which triggered a sharp interest in the studies of the different properties of small-world networks and WS model [7, 8, 9, 10, 11, 12, 13]. The WS model is probably a reasonable illustration of how a small-world network is shaped. However, the small-world effect is much more general, researchers begin to explore other mechanisms producing small-world networks. Recently, Ozik, Hunt and Ott have introduced a simple evolution model (OHO model) of growing small-world networks with geographical attachment preference, where new nodes are linked to geographically nearby ones [14]. A deterministic version of a special case of the OHO model was presented in [15], and further expanded in [16]. In addition, many authors found that small-world properties can be also created in other interesting ways [17, 18].

It is well-known that the WS model is the first successful model that interpolates between a regular ring lattice and a completely random network, and plays an important role in network science. However, the WS model is static (i.e. the network size is fixed), this does not agree with the growth property of many real-life systems [19]. In addition, in real systems, a series of microscopic events shape the network evolution, including addition or removal of a node and addition or removal of

an edge [20, 21, 22, 23, 24]. Therefore, it is interesting to establish a growing small-world model to investigate the effect of local events on the topological features like the WS model. To our best knowledge, all previous models of small-world networks either are static or only took into account the addition of nodes and edges without considering other microscopic events.

In this paper, we present a growing small-world network model controlled by a tunable parameter q , where existing edges can be removed. By tuning parameter q , the model undergoes a phase transition from large to small worlds as the WS model. We study analytically and numerically the structural characteristics, all of which depend on parameter q .

II. THE MODEL

In this section, we introduce a growing model which describes networks from regular to small-world. The model is constructed in the following way (see Fig. 1).

(i) *Initial condition*: We start from an initial state ($t = 2$) of three nodes distributed on a ring, all of which form a triangle.

(ii) *Growth*: At each increment of time, a new node is added which is placed in a randomly chosen internode interval along the ring. Then we perform the following two operations.

(iii) *Addition of edges*: The new node is connected its two nearest nodes (one on either side) previously existing. Nearest, in this case, refers to the number of intervals along the ring.

(iv) *Removal of an edge*: With probability q , we remove the edge linking the two nearest neighbors of the new node.

The growing processes are repeated until the network reaches the desired size.

When $q = 1$, the network is reduced to the one-dimension ring lattice. For $q = 0$, no edge are deleted, the model coincides with a special case of the OHO model [14]. Varying q in the interval (0,1) allows one to study the crossover between the one-dimension regular lattice and the small-world networks.

By construction, at every step, the number of nodes

*Electronic address: zhangzz@fudan.edu.cn

†Electronic address: sgzhou@fudan.edu.cn;
URL: <http://www.iipl.fudan.edu.cn/~zhousg/sgzhou.htm>

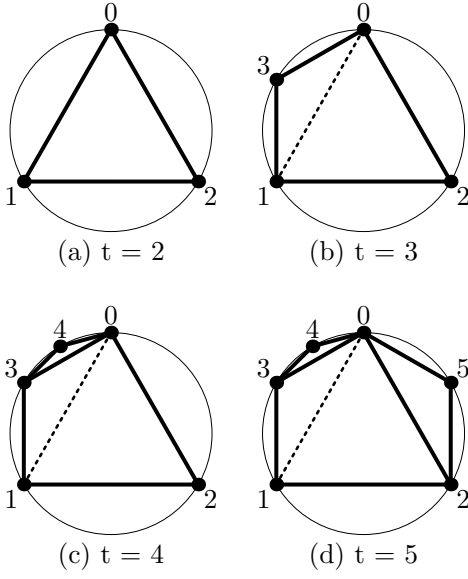


FIG. 1: Scheme of the growing network for the case of $q = 0.3$, showing the first four time steps of the evolving process. The dashed line represents the edge removed.

increases one, while the average number of edges added is $2 - q$. Then we can see easily at time t , the network consists of $t + 1$ nodes and average $(2 - q)t + 2q - 1$ edges. Thus when t is large, the average node degree at time t is equal approximately to a constant value $4 - 2q$, which shows our network is sparse like many real-life networks.

III. TOPOLOGICAL PROPERTIES

We focus on the behavior of the topological characteristics, in terms of the degree distribution, the clustering coefficient, and the average path length, as a function of the parameter q .

A. Degree Distribution

The degree distribution is one of the most important statistical characteristics of a network. For $q = 1$, all nodes have the same number of connections 2, the network exhibits a completely homogeneous degree distribution. Next we focus the case $0 \leq q < 1$. In order to conveniently describe the computation of the network characteristics, we label nodes by their birth times, $s = 0, 1, 2, \dots, t$, and use $p(k, s, t)$ to denote the probability that at time t a node created at time s has a degree k . At time t , there are $t + 1$ internode intervals along the ring and each node has two intervals (one on either side). The master equation [25, 26] governing the evolution of the degree distribution of an individual node has

the form

$$p(k, s, t + 1) = \frac{2(1 - q)}{t + 1} p(k - 1, s, t) + \left(1 - \frac{2(1 - q)}{t + 1}\right) p(k, s, t) \quad (1)$$

with the initial condition, $p(k, s = 0, 1, 2, t = 2) = \delta_{k,2}$ and the boundary one $p(k, t, t) = \delta_{k,2}$. This accounts for two possibilities for a node: first, with probability $\frac{2(1 - q)}{t + 1}$, it may get an extra edge from the new node while its existing edges remain undeleted, and thus increase its own degree by 1; and second, with the complimentary probability $1 - \frac{2(1 - q)}{t + 1}$, the nodes may remain in the former state with the former degree. It should be noted that Eq. (1) and all the following ones are exact for all $t \geq 2$.

The total degree distribution of the entire network can be obtained as

$$P(k, t) = \frac{1}{t + 1} \sum_{s=0}^t p(k, s, t) \quad (2)$$

Using this and applying $\sum_{s=0}^t$ to both sides of Eq. (1), we get the following master equation for the degree distribution:

$$(t + 2)P(k, t + 1) - (t + 1)P(k, t) = 2(1 - q)P(k - 1, t) - 2(1 - q)P(k, t) + \delta_{k,2}. \quad (3)$$

The corresponding stationary equation, i.e., at $t \rightarrow \infty$, takes the form

$$(3 - 2q)P(k) - (2 - 2q)P(k - 1) = \delta_{k,2}. \quad (4)$$

Eq. (4) implies that $P(k)$ is the solution of the recursive equation

$$P(k) = \begin{cases} \frac{2 - 2q}{3 - 2q} P(k - 1) & \text{for } k > 2 \\ 1/(3 - 2q) & \text{for } k = 2 \end{cases} \quad (5)$$

giving

$$P(k) = \frac{1}{3 - 2q} \left(\frac{2 - 2q}{3 - 2q} \right)^{k-2} (k \geq 2), \quad (6)$$

which decays exponentially with k . For $q = 0$, Eq. (6) recovers the result previously obtained in [14]. Thus the resulting network is an exponential network. Note that most small-world networks including the WS model belong to this class [9, 14, 15, 16].

In Fig. 2, we report the simulation results of the degree distribution for several values of q . From Fig. 2, we can see that the degree spectrum of the networks is continuous and the degree distribution decays exponentially for large degree values, in agreement with the analytical results and supporting a relatively homogeneous topology similar to most small-world networks [9, 14, 15, 16].

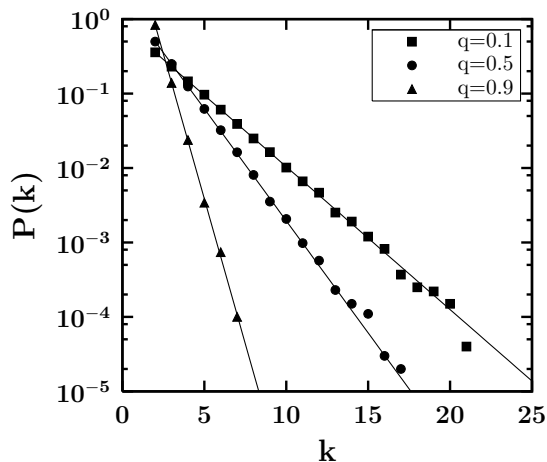


FIG. 2: Semilogarithmic graph of degree distribution of the networks with order $N = 10^5$. The solid lines are the analytic calculation values given by Eq. (6).

B. Clustering coefficient

Most real-life networks show a cluster structure which can be quantified by the clustering coefficient [1, 2, 3, 4, 5]. The clustering coefficient of a node gives the relation of connections of the neighborhood nodes connected to it. By definition, clustering coefficient C_i of a node i is the ratio of the total number e_i of existing edges between all k_i its nearest neighbors and the number $k_i(k_i - 1)/2$ of all possible edges between them, i.e. $C_i = 2e_i/[k_i(k_i - 1)]$. The clustering coefficient C of the whole network is the average of all individual C_i 's.

For the case of $q = 1$, the network is a one-dimensional chain, the clustering coefficient of an arbitrary node and their average value are both zero.

For the case of $q = 0$, using the connection rules, it is straightforward to calculate exactly the clustering coefficient of an arbitrary node and the average value for the network. When a node i enters the network, k_i and e_i are 2 and 1, respectively. After that, if the degree k_i increases by one, then its new neighbor must connect one of its existing neighbors, i.e. e_i increases by one at the same time. Therefore, e_i is equal to $k_i - 1$ for all nodes at all time steps. So there exists a one-to-one correspondence between the degree of a node and its clustering. For a node v with degree k , the exact expression for its clustering coefficient is $\frac{2}{k}$, which has been also been obtained in other models [14, 15, 16, 27, 28]. This expression for the local clustering shows the same inverse proportionality with the degree as those observed in a variety of real-life networks [29]. In this limiting case, the clustering coefficient C of the whole network is given by

$$C = 2 \sum_{k=2}^{\infty} \frac{1}{k} P(k) = \frac{3}{2} \ln 3 - 1 \approx 0.6479. \quad (7)$$

So in the limit of large t the clustering coefficient is very

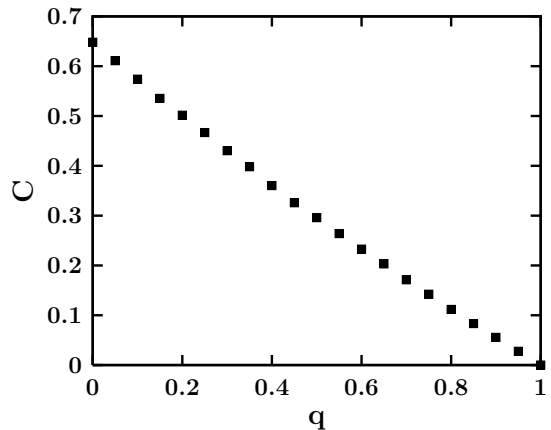


FIG. 3: The clustering coefficient C of the whole network as a function of q .

high.

In the range $0 < q < 1$, it is difficult to derive an analytical expression for the clustering coefficient either for an arbitrary node or for the average of them. In order to obtain the result of the clustering coefficient C of the whole network, we have performed extensive numerical simulations for the full range of q between 0 and 1. Simulations were performed for system sizes 10^5 , averaging over 20 network samples for each value of q .

In Fig. 3, we plot the clustering coefficient C as a function of q . It is obvious that C decreases continuously with increasing q . As q increases from 0 to 1, C drops almost linearly from 0.6479 to 0. Note that although the clustering coefficient C changes linearly for all q , we will show below that in the large limit of q , the average path length changes exponentially as q . This is little different from the phenomenon observed in the WS model where C remains practically unchanged in the process of the network transition to a small world.

C. Average Path Length

Certainly, the most important property of a small-world network is a logarithmic average path length (APL) (with the number of nodes). Here APL means the minimum number of edges connecting a pair of nodes, averaged over all pairs of nodes. It has obvious implications for the dynamics of processes taking place on networks. Therefore, its study has attracted much attention.

For the case of $q = 1$, the average path length increases linearly with network size. For the case of $q = 0$, the network grows stochastically. Generally speaking, for a randomly growing network, the analytical calculation for APL is difficult. Below, we will give an upper bound for the APL of this particular case, which shows that the APL increases at most logarithmically with network size.

If $d(i, j)$ denotes the distance between nodes i and j ,

we introduce the total distance of the network with size N as $\sigma(N)$:

$$\sigma(N) = \sum_{0 \leq i < j \leq N-1} d(i, j), \quad (8)$$

and we denote the APL by $L(N)$, defined as:

$$L(N) = \frac{2\sigma(N)}{N(N-1)}. \quad (9)$$

In the limiting case $q = 0$, the distances between existing node pairs will not be affected by the addition of new nodes. Then we have the following equation:

$$\sigma(N+1) = \sigma(N) + \sum_{i=0}^{N-1} d(i, N). \quad (10)$$

Assume that the node N is added and connected to two nodes w_1, w_2 linked by edge \mathbb{E} , then Eq. (10) can be rewritten as:

$$\begin{aligned} \sigma(N+1) &= \sigma(N) + \sum_{i=0}^{N-1} [D(i, w) + 1] \\ &= \sigma(N) + N + \sum_{i=0}^{N-1} D(i, w), \end{aligned} \quad (11)$$

where $D(i, w) = \min\{d(i, w_1), d(i, w_2)\}$. Constricting the edge \mathbb{E} continuously into a single vertex w (here we assume that $w \equiv w_1$), we have $D(i, w) = d(i, w)$. Since $d(w_1, w) = d(w_2, w) = 0$, Eq. (11) can be rewritten as:

$$\sigma(N+1) = \sigma(N) + N + \sum_{i \in \Gamma} d(i, w), \quad (12)$$

where $\Gamma = \{0, 1, 2, \dots, N-1\} - \{w_1, w_2\}$ is a node set with cardinality $N-2$. The sum $\sum_{i \in \Gamma} d(i, w)$ can be considered as the total distance from one node w to all the other nodes in the network with size $N-1$, which can be roughly evaluated by mean-field approximation in terms of $L(N-1)$ as [30, 31, 32]:

$$\sum_{i \in \Gamma} d(i, w) \approx (N-2)L(N-1). \quad (13)$$

Note that, as $L(N)$ increases monotonously with N , it is clear that:

$$(N-2)L(N-1) = \frac{2\sigma(N-1)}{N-1} < \frac{2\sigma(N)}{N}. \quad (14)$$

Combining Eqs. (12), (13) and (14), one can obtain the inequation:

$$\sigma(N+1) < \sigma(N) + N + \frac{2\sigma(N)}{N}. \quad (15)$$

Considering Eq. (15) as an equation and not an inequality, we can provide an upper for the variation of $\sigma(N)$ as

$$\frac{d\sigma(N)}{dN} = N + \frac{2\sigma(N)}{N}, \quad (16)$$

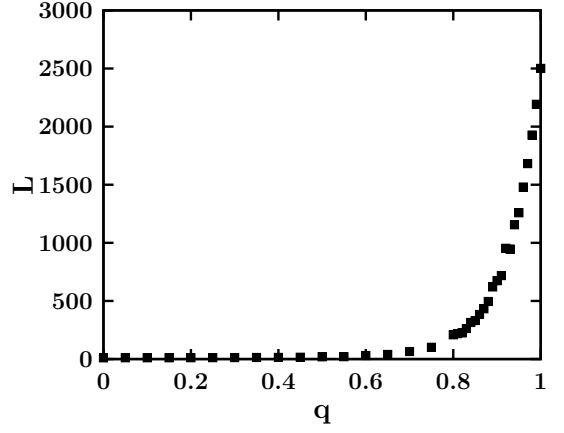


FIG. 4: Average path length L versus parameter q .

which leads to

$$\sigma(N) = N^2(\ln N + \alpha), \quad (17)$$

where α is a constant. As $\sigma(N) \sim N^2 \ln N$, we have $L(N) \sim \ln N$. Note that as we have deduced Eq. (17) from an inequality, then $L(N)$ increases at most as $\ln N$ with N . Therefore, we have proved that in the special case of $q = 0$, there is a slow growth of APL with network size N .

For $0 < q < 1$, in order to obtain the variation of the average path length with the parameter q , we have performed extensive numerical simulations for different q between 0 and 1. Simulations were performed for system sizes 10^4 , averaging over 20 network samples for each value of q . In Fig. 4, we plot the average path length L as a function of q . We observe that, when lessening q from 1 to 0, average path length L drops drastically from a very high value a small one, which predicts that a phase transition from large-world to small-world occurs. This behavior is similar to that in the WS model.

Why is the average path length L low for small q ? The explanation is as follows. The older nodes that had once been nearest neighbors along the ring are pushed apart as new nodes are positioned in the interval between them. From Fig. 1 we can see that when new nodes enter into the networks, the original nodes are not near but, rather, have many newer nodes inserted between them. When q is small, the network growth creates enough "shortcuts" (i.e. long-range edges) attached to old nodes, which join remote nodes along the ring one another as in the WS model [6]. These shortcuts drastically reduces the average path length, leading to a small-world behavior.

IV. CONCLUSIONS

In summary, we have proposed a one-parameter model of growing small-world networks. In our model, in addition to new edges connecting new nodes and old ones,

edges between old nodes may be removed. The presented model interpolates between one-dimensional regular ring and small-world networks, which allow us to explore the crossover between the two limiting cases. We have obtained both analytically and numerically the solution for relevant parameters of the network and observed that our model exhibits the classical phenomenon as that in the WS model. Our model may provide a useful tool to investigate the influence of the clustering coefficient or average path length in different dynamics processes taking place on networks. In addition, using the idea presented here,

one can also construct models interpolating between homogeneous and heterogeneous networks [33].

Acknowledgments

This research was supported by the National Natural Science Foundation of China under Grant Nos. 60373019, 60573183, and 90612007.

-
- [1] M. E. J. Newman, *J. Stat. Phys.* **101**, 819 (2000).
 - [2] R. Albert and A.-L. Barabási, *Rev. Mod. Phys.* **74**, 47 (2002).
 - [3] S. N. Dorogovtsev and J. F. F. Mendes, *Adv. Phys.* **51**, 1079 (2002).
 - [4] M. E. J. Newman, *SIAM Rev.* **45**, 167 (2003).
 - [5] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez and D.-U. Hwang, *Phys. Rep.* **424**, 175 (2006).
 - [6] D. J. Watts and S. H. Strogatz, *Nature* **393**, 440 (1998).
 - [7] M. Barthélemy and L. A. N. Amaral, *Phys. Rev. Lett.* **82**, 3180 (1999).
 - [8] M. E. J. Newman and D. J. Watts, *Phys. Lett. A* **263**, 341 (1999).
 - [9] A. Barrat, and M. Weigt, *Eur. Phys. J. B* **13**, 547 (2000).
 - [10] L. A. N. Amaral, A. Scala, M. Barthélemy and H.E. Stanley, *Proc. Natl. Acad. Sci. U.S.A.*, **97**, 11149 (2000).
 - [11] J. Kleinberg, *Nature* **406**, 845 (2000).
 - [12] F. Comellas, J. Ozón, and J.G. Peters, *Inf. Process. Lett.* **76**, 83 (2000).
 - [13] F. Comellas and M. Sampels, *Physica A* **309**, 231 (2002).
 - [14] J. Ozik, B.-R. Hunt, and E. Ott, *Phys. Rev. E* **69**, 026108 (2004).
 - [15] Z. Z. Zhang, L. L. Rong, and C. H. Guo, *Physica A* **363**, 567 (2006).
 - [16] Z. Z. Zhang, L. L. Rong, and F. Comellas, *J. Phys. A* **39**, 3253 (2006).
 - [17] P. Blanchard and T. Krueger, *Phys. Rev. E* **69**, 046136 (2005).
 - [18] L. H. Yang and M. D. Holland, *Phys. Rev. E* **72**, 067101 (2005).
 - [19] A.-L. Barabási and R. Albert, *Science* **286**, 509 (1999).
 - [20] R. Albert, and A.-L. Barabási, *Phys. Rev. Lett.* **85**, 5234 (2000).
 - [21] S. N. Dorogovtsev, and J. F. F. Mendes, *Europhys. Lett.* **52**, 33 (2000).
 - [22] Q. H. Chen, and D. H. Shi, *Physica A* **335**, 240 (2004).
 - [23] W.-X. Wang, B.-H. Wang, B. Hu, G. Yan, and Q. Ou, *Phys. Rev. Lett.* **94**, 188702 (2005).
 - [24] D. H. Shi, L. M. Liu, X. Zhu, and H. J. Zhou, *Europhys. Lett.* **76**, 10315 (2006).
 - [25] S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin, *Phys. Rev. Lett.* **85**, 4633 (2000).
 - [26] S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin, *Phys. Rev. E* **65**, 062101 (2001).
 - [27] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, *Phys. Rev. E* **65**, 066122 (2002).
 - [28] M. Hinczewski and A. N. Berker, *Phys. Rev. E* **73**, 066126 (2006).
 - [29] E. Ravasz and A.-L. Barabási, *Phys. Rev. E* **67**, 026112 (2003).
 - [30] T. Zhou, G. Yan, and B. H. Wang, *Phys. Rev. E* **71**, 046141 (2005).
 - [31] Z. Z. Zhang, L. L. Rong, and F. Comellas, *Physica A* **364**, 610 (2006).
 - [32] Z. Z. Zhang, L. L. Rong, and S. G. Zhou, *Phys. Rev. E*, **74**, 046105 (2006).
 - [33] J. Gómez-Gardeñes and Y. Moreno, *Phys. Rev. E* **73**, 056124 (2006).